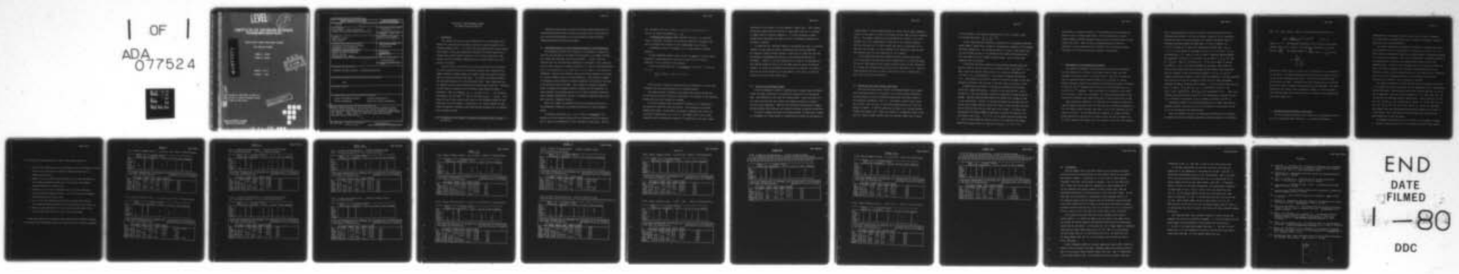
AD-A077 524 CLAREMONT MEN'S COLL CA INST OF DECI ON SCIENCE F/G 14/4
DELTA-EXACT LOWER CONFIDENCE BOUNDS F SERIES SYSTEM RELIABILI--ETC(U)
SEP 79 J B LUCKE , J MYHRE N00014-78-C-0213
UNCLASSIFIED 79-2 NL

1 OF 1
ADA
077524



END
DATE
FILMED
1 -80
DDC

INSTITUTE OF DECISION SCIENCE
FOR BUSINESS & PUBLIC POLICY

DELTA-EXACT LOWER CONFIDENCE BOUNDS

FOR SERIES SYSTEM

JAMES B. LUCKE

JANET M. MYHRE

Report #79-2

August 1979

Research supported in part by
Office of Naval Research grant
N 000 14-78-C-0213

Claremont Men's College
Claremont, California

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 79-2	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Delta-Exact Lower Confidence Bounds for Series System Reliability		5. TYPE OF REPORT & PERIOD COVERED Technical rept.
		6. PERFORMING ORG. REPORT NUMBER 79-2
7. AUTHOR(s) James B. Lucke Janet Myhre		8. CONTRACT OR GRANT NUMBER(s) N00014-78-C-0213
9. PERFORMING ORGANIZATION NAME AND ADDRESS Institute of Decision Science Claremont Men's College Claremont, California 91711		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Reliability
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research 800 N. Quincy Arlington, VA 22217		12. REPORT DATE Sept. 1979
		13. NUMBER OF PAGES 23
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) 25		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

N/A

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Lower Confidence Bounds — Systems Reliability
Steins Estimates — Empirical Bayes Estimates

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Recent modifications of the Likelihood Ratio method of computing lower confidence bounds for series system reliability are shown to have properties necessary for accurate application to actual data. This is shown by comparing various methods using simulations.

DD FORM 1473, 1 JAN 73 — EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

410 359

DELTA-EXACT LOWER CONFIDENCE BOUNDS FOR SERIES SYSTEM RELIABILITY

I. Introduction

This report contains some recent modifications to the Likelihood Ratio Method when computing lower confidence bounds for system reliability from binomial component data. It is shown that these modifications produce bounds which appear to have those properties necessary for accurate application to binomial data usually found in practice, that is from structures of order greater than three binomial component data with unequal sample sizes and a paucity of failures.

Journal articles on comparisons of various confidence bound techniques for series system reliability from binomial component data invariably only compare specific component sample size (usually equal) and specific success data for systems of order two and three. (For example see Mann and Grubbs [8], Myhre and Saunders [11] and Winterbottom [12]). These comparisons are not adequate for data from the majority of systems seen in practice. Therefore, for this report we compare various confidence bound methods by means of simulation. That is, assume $\underline{p} = (p_1, \ldots p_k)$ and $\underline{n} = (n_1, \ldots n_k)$ are given and simulate, say 500 times (1), data of the form $(x_1,n_1),\ldots,(x_k,n_k)$. From these 500 sets of data calculate 500 lower $(1-\alpha)\%$ confidence bounds for each method that is being compared. Comparisons of the various methods are made using the empirical confidence bound distributions obtained from the observed lower confidence bounds.

---

(1) A discussion of the number of simulations performed and why is given on page 10.

Simulation results and the concept of δ-exact bounds demonstrate why the Modified Likelihood Ratio Methods yield bounds which are preferable to those obtained using methods such as Approximately Optimum, Modified Maximum Likelihood and Modified Log Gamma.

II. Desirable Properties of Lower Confidence Bounds on System Reliability

Many expensive systems such as missiles, are produced to complete one task only and they must be certified to the purchaser to be highly reliable even though they often may not be tested directly. In such a case the producer will test the individual components or subsystems of the system and obtain a point estimate of systems reliability, R, from the resulting binomial data: x_i successes out of n_i trials for $1 \leq i \leq k$, where k is the number of components. Further, most contracts require an interval estimate in the form of a lower confidence bound on system reliability. The estimation problem of obtaining a point estimate for system reliability is much easier than that of obtaining an interval estimate. Even though the literature contains a number of methods for computing interval estimates for the reliability of series systems, these methods usually do not produce acceptable bounds for highly reliable systems where the data are disparate component sample sizes and a paucity of component failures.

There are a number of properties that have been used historically in the mathematical literature as desirable properties for lower confidence bounds:

A confidence interval $\underline{R}\,(x_1,\ldots,x_k)$ is said to be <u>unbiased</u> if the probability of the interval covering the true parameter value R is greater than or equal to the probability of its covering any false value. That is,

for confidence level $1-\alpha$, $P\,[\underline{R}\,(x_1,\ldots,x_k) \le R'] \le 1-\alpha$ for all $R' < R$ and for all nuisance parameters $p_1,\ldots,p_k$.

A <u>uniformly most accurate</u> (UMA) confidence interval, in a specified class of intervals, is based on the acceptance region of a hypothesis test that is uniformly most powerful in the corresponding class of tests.

A confidence interval is optimum if it is a UMA unbiased confidence interval.

A lower confidence bound of level $1-\alpha$ for R is <u>exact</u> if the lower limits obtained in random sampling are less than or equal to R on *at least* a proportion $1-\alpha$ of occasions in the long run.

A lower confidence bound for R is <u>consistent</u> if for every $\varepsilon > 0$ and all $1 \le i \le k$

$$\lim_{n_i \to \infty} P\,[\,|\underline{R}\,(x_1,\ldots,x_k) - R| < \varepsilon] = 0$$

There are also a number of properties which are desirable or in fact even necessary for a practical application of confidence bounds:

The producer needs a confidence bound method that produces bounds that are as close to the true reliability as possible: The variance of the confidence bounds about the true R should be a minimum.

It is also necessary that the method be flexible so it is applicable to the variety of cases that arise in practice. The constraints of time and expense often result in data about the system's components that consist of a variety of sample sizes, n_i, $1 \le i \le k$. The sample sizes of all or many components may be constrained by expense or configuration to be small, either

absolutely or with respect to other component's sample sizes. Thus a method should work equally well for data due to equal sample sizes or to a variety of sample sizes on different components. It is also necessary to have a method that is applicable when some or all of the components experience no failures in testing.

It would also be a desirable feature if the method were easy to calculate; however, for most cases the above considerations are far more important. For expensive, highly reliable systems it seems fair to assume that the time and expense of programming and running a more desirable method will be worthwhile. Finally, it is often necessary for a method to be applicable to more complex systems as well as for the simple cases of series or parallel systems. Of course if one method could be shown to be superior in a special case such as series systems but not applicable to other cases, it would be certain to be used in that special case.

III. δ-exact Lower Confidence Bounds

There are presently a variety of methods used to compute lower confidence bounds on system reliability but no method has been shown to be universally best. Some methods will only work for series systems; others will work for all types of systems but have trouble with samples containing no failures for one or more components. Some are too conservative when applied to highly reliable systems while other are not bounds for systems of moderate reliability.

In order to enlarge the field of available methods, in particular to admit two refinements of a known method to consideration we define the new concept of

δ-exact bound: A lower confidence bound is a δ-exact $(1-\alpha)\%$ lower confidence bound if in the long run $(1-\alpha)\%$ of the bounds are less than the true reliability plus δ. For example, assume the true reliability of a system is .92, δ=.01 and 80% l.c.b.'s are computed then at least 80% of the interval estimates would be less than .93. If such a method were used on a particular set of data and a δ=.02, 90% l.c.b. were estimated to be .88 then we could say with 90% certainty that the true reliability is .86 or higher.

In our simulations there are examples in which a method that gives a δ=.01, 90% bound can be much better than a method that is a δ=.00, 90% bound; for the .01, 90% bound may have standard deviation about the true reliability of .07 while the standard deviation of the δ=.00, 90% bound is .14. Thus the added flexibility of being a δ-exact bound may pay off in a method that is much closer to the true reliability.

IV. Methods Used for Series Systems Comparisons

There has been much recent interest in developing methods that compute lower confidence bounds for the general case of series system reliability. An early method for series and parallel systems that has found acceptance was presented by Madansky in 1965 [7]. This method is based on the result, due to Wilks, that $-2\ln L(P_o)$ is distributed asymptotically as χ^2 with one degree of freedom. In 1968 the method was extended to all coherent systems by Myhre and Saunders [10]. While this method, as documented in 1968, is not exact for highly reliable systems where the component sample size is small,

it is applicable to any coherent system and can in fact be shown to work for systems where $H(\underline{p}) = \sum_{j=1}^{m} w_j \left(\prod_{i=1}^{k} P_i^{\alpha_{ij}} \right)$.

In 1972 Easterling [2] introduced a modification to the maximum likelihood method (MMLI), a method which is based on the asymptotic normality of the maximum likelihood estimations. This method was shown to be good in some simple cases but that the MMLI as well as the likelihood ratio method have problems with data that might arise for a highly reliable system: data for which some components have no failures.

In 1974 Mann and Grubb presented the Approximately Optimum method (AO) which is based on the fact that $P\,[\, R \geq \exp\{-m[1-(v/9m^2) + z_\alpha (v/3m)]^3\}] \simeq 1-\alpha$ where m and v are the mean and variance of the posterior distribution of $-\ln R$ and $-\ln R$ is approximated by a non-central chi-square distribution [8]. This method is restricted to series (or parallel) systems but has seen extensive use owing to the fact that it has been shown to be close to exact.

In 1968 Borsting and Woods [1] introduced a method based on fitting $-\log \hat{R}$ with a gamma distribution. Initially it was used only for series systems but it has been modified by Woods for application to some non-series systems. Recently Tomsky further modified the method in an attempt to make it less conservative. However, in testing this method it was found that for a given set of sample sizes the method often gives a lower estimate from data with no failures than from data with one or two failures. To correct this inconsistency the method (MLGB) was modified to estimate the bound for the zero failure case ($x_i = n_i$, $1 \leq i \leq k$) as .01 larger than the estimate for the case in which exactly one failure occurs on the component with largest sample size ($x_j = n_j - 1$ where $n_j = \max_i \{n_i\}$ and $x_i = n_i$, $1 \leq i \leq k$, $i \neq j$).

This method is of special interest as it was designed for the practical but difficult case of unequal sample sizes. The above methods are a fairly complete sample of methods presently known to work for series systems of order greater than three. There is also the weighting method [9] which is applicable to series systems of high reliability but this method requires knowledge of the relative ordering of component unreliability.

V. Improvements in the Likelihood Ratio Method

As originally introduced the Likelihood Ratio method had difficulties when the system was so reliable or the samples taken so small that some components are tested without a resulting failure. Since this is often the case for expensive highly reliable systems such as missiles, an initial attempt was made to correct this difficulty by adjusting the number of failures to $4/kn_i$ for the i^{th} component when there were no actual failures out of n_i tests for the i^{th} component. Thus this factor will be smaller as the order of the system, k, and/or as n_i increases. The adjusted binomial data is then used by the likelihood ratio method resulting in a method referred to as LRL.

When testing the components it is known that no component is perfect but time and expense constraints limit the test size of the different components so that for highly reliable systems, some, many or all of the components may not register a failure. On the other hand if a component in a highly reliable system registers one failure out of fifteen trials, the lack of failures on the other components may suggest that the component is more than .933 reliable.

Thus it seems desirable to adjust the failures (and successes) used by the likelihood ratio method to reflect the overall reliability of the components. A method that will do such an adjustment for normally distributed data, by shrinking everything towards the grand mean, is Stein's estimator [5]. In [6] it is shown how to adapt this technique to binomial data: Let x_i, $1 \le i \le k$, be the number of successes out of n_i, $1 \le i \le k$ trials; let $y_i = n_i \arcsin(2x_i/n_i) + 1)$, where $y_i \sim N(\sigma_i, 1)$. Stein's emperical Bayesian method can be applied so that $\hat{y}_i = \bar{y} + (1-(k-3)/s)\ (y_i - \bar{y})$ where $s = \sum_1^k (y_i - \bar{y})^2$ and $\bar{y} = (\sum_1^k y_i)/k$; then $\hat{r}_i = .5(1+\sin \hat{y}_i/\sqrt{n_i})$ is the new estimate of success rate and $\hat{x}_i = \hat{r}_i n_i$ is the Stein's Estimate of successes. This method was tried but the results were often contradictory with failure rates going away from the "average" rather than shrinking towards it. In [3], [4] and [6] "limited translations estimators" were defined and explained. They were appealing since they do not allow the estimates to shift "too far" from the maximum likelihood estimate. Using the above notation: $\hat{y}_i = \bar{y} + (1-\{(k-3)/v\}\phi(u_i))(y_i - \bar{y})$, where $v = \max\{n, \sum_1^k (y_i - y)^2\}$ and $u_i = (\ (y_i - \bar{y})^2 k(k-3)\ )/((k-1)v)$ and $\phi(u_i) = \min\{1, D_{80}(k-1)/\sqrt{u_i}\}$ where D_{80} is the value given in [4]. After this "limited translation Steins estimate" of the binomial data is complete it used by the likelihood ratio method to get an estimate, referred to as LRS 13.

Preliminary testing of this method indicated that it might have some very desirable features even being a δ=.01 or δ=.02 bound it was noticeably closer to the true reliability than other methods.

Since the component test data is binomially distributed and not normally distributed an empirical Bayesian approach was directly applied to the binomial

data. Let $X_i | p \sim B(n_i, p)$ where p is distributed as Beta with p.d.f.

$$g(p) = \frac{\Gamma(n)}{\Gamma(r)\Gamma(n-r)} p^{r-1} (1-p)^{n-r-1} \qquad 0 < p < 1$$

where $n > r > 0$. The marginal distribution of X_i is used to obtain point estimates for r and n, $\tilde{r}$ and $\tilde{n}$. The posterior distribution of $p|x_i$ has $E[p|x_i] = \frac{r+x_i}{n+n_i}$. We estimate $E[p|x_i]$ by $\frac{\tilde{r}+x_i}{\tilde{n}+n_i}$, where $\tilde{r} = \sum_{i=1}^{k} x_i/k$ and

$$\tilde{n} = \frac{\sum_{i=1}^{k} x_i}{\sum_{i=1}^{k} \left(\frac{x_i}{n_i}\right)}$$

Note that it follows that if $n_i = N$ for all i then $\tilde{n} = N$ and $\tilde{r}$ is the average of the successes. The resulting adjusted empirical Bayesian binomial test data is used in the Likelihood Ratio Method. The resulting confidence bounds are referred to as LRB 9. Both LRS 13 and LRB 9 methods had to be adjusted for the case when no failures occurred for any of the components. For these methods the zero failure case is computed by adding .01 to the confidence bound obtained when only one failure occurs, a failure on the component with the largest sample size.

VI. Simulation Model and Methods of Comparison

As all the methods presented, both old and new, are approximate there is no way to analytically show superiority. To establish advantages or

disadvantages of the methods relative to each other they must be compared on a large number of systems with variety in number of components, component sample sizes and system reliability.

To do these comparisons the methods were applied to simulated data from a series system of order k, with component i being tested n_i times, $1 \le i \le k$. Each system was simulated 100 times since a larger number of simulations resulted in the same type of results. The use of 100 simulations is checked periodically by simulating 200 to 500 times and checking these results with the 100 simulation case. For each of the 100 simulations for a method we counted the number of times the computed lower confidence bounds were below the true reliability, below the true reliability plus .01, and below the true reliability plus .02; also the standard deviation of the estimates about the true reliability was computed to give a measure of closeness.

The simulations were run for 80%, 90% and 95% lower confidence bounds but in the interest of brevity the results will be summarized only for the important case of 90% bounds. The simulations were run for systems with k=6, 10, 15, 20 and 30 components, with true system reliability R = .85, .868, .886, .90, .93, .96 and with the following component sample sizes: equal n_i with n_i= 15, 20, 30 or 40; unequal n_i with n_i taking on the values: (5, 10, 20, 30), (15, 30, 60), (20, 30, 40, 50), (20, 20, 20, 20, 100) or (5, 20, 20, 20, 100) where these sample sizes are distributed proportionally according to the order, k, of the system. This covers many of the cases that arise in practice and gives enough data to see the trends.

To summarize the results for a set of cases it is necessary to compare each method to the others and to try to determine an overall "best" method.

In the tables that follow method A is said to have beaten method B if

1. Both are exact 90% bounds (δ=.01, 90% bounds, δ=.02, 90% bounds, or neither result is a bound) and the estimated standard deviation of A is smaller than that of B.
2. Method A is a δ=.01, 90% bound and Method B is an exact 90% bound but the estimated standard deviation of A is less than the estimated standard deviation of B minus .005.
3. Similarly if A is a δ=.02, 90% bound and B is a δ=.01, 90% bound but the estimated standard deviation of A is less than the estimated standard deviation of B minus .005.
4. A is a δ=.02, 90% confidence bound and B is an exact 90% bound but the estimated standard deviation of A is less than that of B minus .01.
5. Method A cannot beat Method B if A is neither an exact 90%, a δ=.01 or δ=.02, 90% bound and B is at least a δ=.02, 90% bound.

Using this criterion the following tables show how each method did relative to the others and which methods seem best overall for the set of cases considered.

TABLE I

k= 6, Equal Sample Sizes , n_i = 15 or 20, 9 Sets of Simulations

	best overall	δ-exact bound δ=0 (exact)	δ=.01	δ=.02	not a bound
AO		9			
MMLI		8			1
MLGB		8			1
LRL	1	5	3		1
LRS13	8	8			1
LRB9		8			1

Pairwise Comparison of Methods for the Sets of Simulations	AO	MMLI	MLGB	LRL	LRS13	LRB9	Comparison of Individual Methods with all Others
AO	–	9-0	2-7	1-8	1-8	1-8	1-4 1-4
MMLI	0-9	–	0-9	2-7	0-9	0-9	0-5
MLGB	7-2	9-0	–	8-1	0-9	3-6	3-2
LRL	8-1	7-2	1-8	–	0-9	2-7	2-3
LRS13	8-1	9-0	9-0	9-0	–	8-1	5-0
LRB9	8-1	9-0	6-3	7-2	1-8	–	4-1

k= 6, Equal Sample Sizes , n_i = 30 or 40, 8 Sets of Simulations

	best overall	δ-exact bound δ=0 (exact)	δ=.01	δ=.02	not a bound
AO		8			
MMLI		8			
MLGB		6	1	1	
LRL		6	1	1	
LRS13	7	4		3	1
LRB9	1	7	1		

Pairwise Comparison of Methods for the Sets of Simulations	AO	MMLI	MLGB	LRL	LRS13	LRB9	Comparison of Individual Methods with all Others
AO	–	8-0	5-3	7-1	1-7	4-4	3 1/2 -1 1/2
MMLI	0-8	–	0-8	0-8	1-7	0-8	0-5
MLGB	3-5	8-0	–	2-6	1-7	1-7	1-4
LRL	1-7	8-0	6-2	–	1-7	2-6	2-3
LRS13	7-1	7-1	7-1	7-1	–	7-1	5-0
LRB9	4-4	8-0	7-1	6-2	1-7	–	3 1/2- 1 1/2

TABLE II

k= 6 , 14 Sets of Simulations, Unequal Sample Sizes
n_i= 20,20,20,20,20,100 or n_i= 5,20,20,20,20,100

	best overall	δ-exact bound δ=0 (exact)	δ=.01	δ=.02	not a bound
AO	3	14			
MMLI				2	6
MLGB	2	2	1	4	7
LRL		10			4
LRS13	1	6	2	1	5
LRB9	8	7	2	1	4

Pairwise Comparison of Methods for the Sets of Simulations	AO	MMLI	MLGB	LRL	LRS13	LRB9	Comparison of Individual Methods with all Others
AO		12-2	8-6	2-12	3-11	2-12	2-3
MMLI	2-12	-	0-14	3-11	2-12	0-14	0-5
MLGB	6-8	14-0	-	8-6	2-12	2-12	2-3
LRL	12-2	11-3	6-8	-	1-13	0-14	2-3
LRS13	11-3	12-2	12-2	13-1	-	1-13	4-1
LRB9	12-2	14-0	12-2	14-0	13-1	-	5-0

k= 6 , 9 Sets of Simulations, Unequal Sample Sizes
n_i= 15,30,60,15,30,60 or n_i= 20,30,40,50,20,30,

	best overall	δ-exact bound δ=0 (exact)	δ=.01	δ=.02	not a bound
AO		9			
MMLI		4	2	2	1
MLGB	1	1	2	3	3
LRL		6	2	1	
LRS13	4	5	1	2	1
LRB9	4	6	1	2	

Pairwise Comparison of Methods for the Sets of Simulations	AO	MMLI	MLGB	LRL	LRS13	LRB9	Comparison of Individual Methods with all Others
AO	-	1-8	4-5	0-9	1-8	0-9	0-5
MMLI		-	3-6	2-7	0-9	0-9	1-4
MLGB	5-4	6-3	-	3-6	2-7	2-7	2-3
LRL	9-0	7-2	6-3	-	1-8	0-9	2-3
LRS13	8-1	9-0	7-2	8-1	-	5-4	5-4
LRB9	9-0	9-0	7-2	9-0	4-5	-	4-1

TABLE III

k= 6 , 9 Sets of Simulations, Unequal Sample Sizes
n_i= 15,15,15,15,15,75 or n_i= 5,15,15,15,15,75

	best overall	δ-exact bound δ=0 (exact)	δ=.01	δ=.02	not a bound
AO	1	9			
MMLI		4			5
MLGB	1	2	2		5
LRL	1	7	1		1
LRS13	1	6	1		2
LRB9	5	7			2

Pairwise Comparison of Methods for the Sets of Simulations	AO	MMLI	MLGB	LRL	LRS13	LRB9	Comparison of Individual Methods with all Others
AO	-	6-3	5-4	1-8	2-7	2-7	2-3
MMLI	3-6	-	0-9	2-7	0-9	0-9	0-5
MLGB	4-5	9-0	-	5-4	3-6	2-7	2-3
LRL	8-1	7-2	4-5	-	1-8	1-8	2-3
LRS13	7-2	9-0	6-3	8-1	-	3-6	4-1
LRB9	7-2	9-0	7-2	8-1	6-3	-	5-0

k= 6 , 5 Sets of Simulations, Unequal Sample Sizes
n_i= 5,10,20,30,5,10

	best overall	δ-exact bound δ=0 (exact)	δ=.01	δ=.02	not a bound
AO		5			
MMLI		4			
MLGB	4	2		2	1
LRL		5			
LRS13	1	5			
LRB9		5			

Pairwise Comparison of Methods for the Sets of Simulations	AO	MMLI	MLGB	LRL	LRS13	LRB9	Comparison of Individual Methods with all Others
AO	-	1-4	1-4	0-5	0-5	0-5	0-5
MMLI	4-1	-	1-4	4-1	2-3	2-3	2-3
MLGB	4-1	4-1	-	4-1	4-1	4-1	5-0
LRL	5-0	1-4	1-4	-	0-5	0-5	1-4
LRS13	5-0	3-2	1-4	5-0	-	0-5	3-2
LRB9	5-0	3-2	1-4	5-0	5-0	-	4-1

TABLE IV

$k=10$, Equal Sample Sizes, $n_i=15$ or 20, 17 Sets of Simulations

	best overall	δ-exact bound $\delta=0$ (exact)	$\delta=.01$	$\delta=.02$	not a bound
AO	2	16		1	
MMLI		17			
MLGB		11	4		2
LRL		10	2		5
LRS13	11	7	3	2	5
LRB9	4	15		2	

	Pairwise Comparison of Methods for the Sets of Simulations						Comparison of Individual Methods with all Others
	AO	MMLI	MLGB	LRL	LRS13	LRB9	
AO	-	17-0	4-13	10-7	6-11	2-15	2-3
MMLI	0-17	-	2-15	4-13	5-12	0-17	0-5
MLGB	13-4	15-2	-	11-6	3-14	3-14	3-2
LRL	7-10	13-4	6-11	-	5-12	4-13	1-4
LRS13	11-6	12-5	14-3	12-5	-	12-5	5-0
LRB9	15-2	17-0	14-3	13-4	5-12	-	4-1

$k=10$, Equal Sample Sizes, $n_i=30$ or 40, 15 Sets of Simulations

	best overall	δ-exact bound $\delta=0$ (exact)	$\delta=.01$	$\delta=.02$	not a bound
AO	9	12	3		
MMLI		15			
MLGB		9	2	3	1
LRL		9	3	2	1
LRS13	5	2		3	10
LRB9	1	9	2	3	1

	Pairwise Comparison of Methods for the Sets of Simulations						Comparison of Individual Methods with all Others
	AO	MMLI	MLGB	LRL	LRS13	LRB9	
AO	-	15-0	12-3	14-1	10-5	13-2	4-1
MMLI	0-15	-	4-11	1-14	10-5	1-14	1-4
MLGB	3-12	11-4	-	3-12	9-6	2-13	2-3
LRL	1-14	14-1	12-3	-	10-5	5-10	3-2
LRS13	5-10	5-10	6-9	5-10	-	6-9	0-5
LRB9	2-13	14-1	13-2	10-5	9-6	-	4-1

TABLE V

k= 10, 6 Sets of Simulations, Unequal Sample Sizes
n_i= 5,10,20,30,5,10,20,30

	best overall	δ-exact bound δ=0 (exact)	δ=.01	δ=.02	not a bound
AO		6			
MMLI		3			
MLGB	3	2	1		3
LRL		6			
LRS13	1	6			
LRB9	1	6			

Pairwise Comparison of Methods for the Sets of Simulations	AO	MMLI	MLGB	LRL	LRS13	LRB9	Comparison of Individual Methods with all Others
AO	-	0-6	3-3	1-5	0-6	0-6	1/2-4 1/2
MMLI	6-0	-	2-4	3-3	2-4	2-4	1 1/2-3 1/2
MLGB	3-3	4-2	-	3-3	3-3	3-3	3-2
LRL	1-5	3-3	3-3	-	0-6	0-6	1-4
LRS13	0-6	4-2	3-3	6-0	-	1-5	1 1/2-3 1/2
LRB9	6-0	4-2	3-3	6-0	5-1	-	4 1/2- 1/2

k= 10, 12 Sets of Simulations, Unequal Sample Sizes
n_i=15,30,60,15,30,60,15,30,60,15 or n_i=20,30,40,50,20,30,40,50,20,30

	best overall	δ-exact bound δ=0 (exact)	δ=.01	δ=.02	not a bound
AO		12			
MMLI		7	2	1	2
MLGB	1	4	3	3	2
LRL		4	3	4	1
LRS13	9	9	2	1	
LRB9	2	10	1	1	

Pairwise Comparison of Methods for the Sets of Simulations	AO	MMLI	MLGB	LRL	LRS13	LRB9	Comparison of Individual Methods with all Others
AO	-	2-10	2-10	2-10	1-11	0-12	0-5
MMLI	10-2	-	3-9	0-12	0-12	0-12	1-4
MLGB	10-2	9-3	-	4-8	2-10	3-9	2-3
LRL	10-2	12-0	8-4	-	1-11	1-11	3-3
LRS13	11-1	12-0	10-2	11-1	-	11-1	5-0
LRB9	12-0	12-0	9-3	11-1	1-11	-	4-1

TABLE VI

k=15, Equal Sample Sizes , n_i=15 or 20 ,10Sets of Simulations

	best overall	δ-exact bound δ=0 (exact)	δ=.01	δ=.02	not a bound
AO	1	9	1	-	-
MMLI	-	10	-	-	-
MLGB	-	8	1	-	1
LRL	-	4	5	-	1
LRS13	7	6	1	1	2
LRB9	2	9	-	1	-

Pairwise Comparison of Methods for the Sets of Simulations	AO	MMLI	MLGB	LRL	LRS13	LRB9	Comparison of Individual Methods with all Others
AO	-	10-1	3-7	6-4	3-7	1-9	2-3
MMLI	0-10	-	1-9	1-9	3-7	0-10	0-5
MLGB	7-3	9-1	-	8-2	2-8	0-10	3-2
LRL	4-6	9-1	2-8	-	2-8	0-10	1-4
LRS13	7-3	7-3	8-2	8-2	-	7-3	5-0
LRB9	9-1	10-1	10-0	10-0	3-7	-	4-1

k=15, Equal Sample Sizes , n_i=30 or 40 , Sets of Simulations

	best overall	δ-exact bound δ=0 (exact)	δ=.01	δ=.02	not a bound
AO	9	10	-	-	-
MMLI	-	10	-	-	-
MLGB	-	7	1	1	1
LRL	-	7	1	1	1
LRS13	1	-	-	1	9
LRB9	-	7	2	-	1

Pairwise Comparison of Methods for the Sets of Simulations	AO	MMLI	MLGB	LRL	LRS13	LRB9	Comparison of Individual Methods with all Others
AO	-	10-0	10-0	10-0	9-1	9-1	5-0
MMLI	0-10	-	2-8	1-9	9-1	1-9	1-4
MLGB	0-10	8-2	-	1-9	8-2	0-10	2-3
LRL	0-10	9-1	9-1	-	9-1	2-8	3-2
LRS13	1-9	2-8	2-8	1-9	-	2-8	0-5
LRB9		10-0	10-0	8-2	8-2	-	4-1

TABLE VII

k=15, 9 Sets of Simulations, Unequal Sample Sizes
n_i = 5,10,20,30,5,10,20,30,5,10,20,30,5,10,20,30,5,10,20,30
or n_i=20,20,20,20,20,100,20,20,20,20,20,100,20,20,20,20,20,100,20,20
or n_i=15,30,60,15,30,60,15,30,60,15,30,60,15,30,60,15,30,60

	best overall	δ-exact bound δ=0 (exact)	δ=.01	δ=.02	not a bound
AO	1	9			
MMLI		4	2		3
MLGB	1	3	1		5
LRL		7	2		
LRS13	5	7	1		1
LRB9	2	7	1		1

Pairwise Comparison of Methods for the Sets of Simulations	AO	MMLI	MLGB	LRL	LRS13	LRB9	Comparison of Individual Methods with all Others
AO	-	5-4	4-5	2-7	1-8	1-8	1-4
MMLI	4-5	-	4-5	2-7	2-7	2-7	0-5
MLGB	5-4	5-4	-	4-5	1-8	1-8	2-3
LRL	7-2	7-2	5-4	-	1-8	1-8	3-2
LRS13	8-1	7-2	8-1	8-1	-	5-4	5-0
LRB9	8-1	7-2	8-1	8-1	4-5	-	4-1

TABLE VIII

$k=20$, Equal Sample Sizes, $n_i = 15$ or 20, 7 Sets of Simulations

	best overall	ϵ-exact bound $\epsilon=0$ (exact)	$\epsilon=.01$	$\epsilon=.02$	not a bound
AO		6	1		
MMLI		7			
MLGB		5	1		1
LRL		4	1	1	1
LRS13	4	3	1		3
LRB9	3	7			

Pairwise Comparison of Methods for the Sets of Simulations	AO	MMLI	MLGB	LRL	LRS13	LRB9	Comparison of Individual Methods with all Others
AO	-	7-0	2-5	3-4	3-4	0-7	1-4
MMLI	0-7	-	1-6	2-5	3-4	0-7	0-5
MLGB	5-2	6-1	-	7-0	1-6	2-5	3-2
LRL	4-3	5-2	0-7	-	2-5	0-7	2-3
LRS13	4-3	4-3	6-1	5-2	-	4-3	5-0
LRB9	7-0	7-0	5-2	7-0	3-4	-	4-1

$k=20$, Equal Sample Sizes, $n_i=30$ or 40, Sets of Simulations

	best overall	ϵ-exact bound $\epsilon=0$ (exact)	$\epsilon=.01$	$\epsilon=.02$	not a bound
AO	4	6	-	-	-
MMLI	-	6	-	-	-
MLGB	-	4	1	-	1
LRL	-	3	2	-	1
LRS13	1	1	1	-	4
LRB9	1	4	-	1	1

Pairwise Comparison of Methods for the Sets of Simulations	AO	MMLI	MLGB	LRL	LRS13	LRB9	Comparison of Individual Methods with all Others
AO	-	6-0	6-0	6-0	5-1	4-2	5-0
MMLI	0-6	-	2-4	1-5	5-1	1-5	1-4
MLGB	0-6	4-2	-	1-5	4-2	0-6	2-3
LRL	0-6	5-1	5-1	-	5-1	1-5	3-2
LRS13	1-5	1-5	2-4	1-5	-	2-4	0-5
LRB9	2-4	5-1	6-0	5-1	4-2	-	4-1

TABLE IX

k= 20, 10 Sets of Simulations, Unequal Sample Sizes
n_i= 5,10,20,30,5,10,20,30,5,10,20,30,5,10,20,30,5,10,20,30
or n_i=20,20,20,20,20,100,20,20,20,20,20,100,20,20,20,20,20,100,20,20
or n_i=15,30,60,15,30,60,15,30,60,15,30,60,15,30,60,15,30,60,15,30

	best overall	δ-exact bound δ=0 (exact)	δ=.01	δ=.02	not a bound
AO	1	10			
MMLI		7	1		2
MLGB	2	3	1	2	4
LRL		7	3		
LRS13	4	8		1	1
LRB9	3	7		2	1

Pairwise Comparison of Methods for the Sets of Simulations	AO	MMLI	MLGB	LRL	LRS13	LRB9	Comparison of Individual Methods with all Others
AO	-	4-6	4-6	1-9	1-9	1-9	0-5
MMLI	6-4	-	3-7	4-6	4-6	1-9	1-4
MLGB	6-4	7-3	-	6-4	3-7	2-8	3-2
LRL	9-1	6-4	4-6	-	1-9	1-9	2-3
LRS13	9-1	6-4	7-3	9-1	-	5-5	4 1/2-1/2
LRB9	9-1	9-1	8-2	9-1	5-5	-	4 1/2-1/2

VII. Conclusions

From the summary data in the above tables we can see some interesting patterns and develop some guidelines for which methods should be used and for what types of systems. In particular it is clear that for many standard, "textbook" situations, the AO method is superior; that is, the tables show that if there are a large number of components in a series system and if all the components have equal, moderate to large, sample sizes, then the approximate properties of AO take hold and it gives the "best" bounds. On the other hand if the situation is complicated by small component sizes or if the component sample sizes are unequal, then the AO method is worse than many of the other methods due to its being too conservative and having the largest estimated standard deviation about the true reliability (in all the simulations AO was at least a δ=.02, 90% bound and in most cases it was a δ=0 bound).

In the series situations that most closely resemble those found in actual practice, i.e. either the sample sizes are small or the sample sizes are unequal the methods based on the Likelihood Ratio method are consistently better than all the others. In particular for 6 or a larger number of components with equal but small sample sizes (n_i= 15 or 20), LRS 13 is consistently the best though often it is a 90% bound with δ=.01 or .02 rather than δ=0. For large sample sizes, LRS 13 is too optimistic and is often not even a δ=.02, 90% bound.

A more consistent method in the equal sample size case is LRB 9 which is second to AO in the equal, not small, component sample size case and second to LRS 13 in the equal, small component sample size case. LRB 9 seldom fails to be a bound losing to LRS 13 only because it is not as close to the true

reliability as LRS 13. When LRB 9 loses to AO it just barely loses.

In the most complicated, bur also most realistic, case where the sample size of the components of the system vary the LRB 9 and LRS 13 methods are consistently superior to the other methods. This is partially due to the fact that the Likelihood Ratio is not as conservative in the unequal sample case as are the other methods. Again LRS 13 sometimes fails to be even a δ=.02 bound, whereas LRB 9 is consistently a bound and in these cases it is as close or closer to the true reliability than LRS 13. For the unequal sample case the only other method that comes close to either LRB 9 or LRS 13 is MLGB which does well only in one type of case: small unequal sample sizes (n_i varies amont 5, 10, 20, 30). MMLI and AO do not do well at all for the unequal sample cases for they are so conservative that they are usually 100% bounds with high estimated standard deviations.

For computing (90%) lower confidence bounds for series systems from binomial test data we can therefore recommend: a. Use the AO method for the equal sample size with moderate to large sample size (n=40 or more). b. Use LRS 13 for equal small sample size case. c. Use LRB 9 in all cases since it is very comparable to the AO in the AO's best cases and is often better than LRS 13 in the unequal sample size case.

References

[1] Borsting, J.R. and Woods, W.M. "A Method for Computing Lower Confidence Limits on System Reliability using Component Failure Data with Unequal Sample Sizes" 1968 - U.S. Naval Postgraduate School.

[2] Easterling, R.G. "Approximate Confidence Limits for System Reliability" JASA 67 (1972) p. 220-222.

[3] Efron, B. and Morris, C. "Limiting the Risk of Bayes Estimations - Part 1: The Bayes Case" JASA, 66 (1971) p. 807-815.

[4] _______________ "Limiting the Risk - Part 2: The Empirical Bayes Case" JASA 67 (1972) p. 130-139.

[5] _______________ "Steins Estimation Rule and its Competitors, An Empirical Approach" JASA 68 (1973) p. 117-130.

[6] _______________ "Data Analysis Using Stein's Estimator" JASA 70 (1975) p. 311-319.

[7] Madansky, A. "Approximate Confidence Limits for the Reliability of Series and Parallel Systems" Technometrics 7 (1965) p. 495-503.

[8] Mann, N.R. and Grubbs, F.E. "Approximately Optimum Confidence Bounds for Systems Reliability Based on Component Test Data" Technometrics 16 No.3 (1974) p. 335-347.

[9] Myhre, J.M., Rosenfeld, A.M., Saunders, S.C. "Determining Confidence Bounds for Highly Reliable Coherent Systems Based on a Paucity of Component Failures," Naval Research Logistics Quarterly 25 (1978) p. 213-227.

[10] Myhre, J.M. and Saunders, S.C. "On Confidence Limits for the Reliability of Systems," Annals of Mathematical Statistics 39 No.5 (1968) p. 1463-1472.

[11] Myhre, J.M. and Saunders, S.C., "Comparison of Two Methods of obtaining Approximate Confidence Intervals for System Reliability", Technometrics 10 No. 1 (February 1968) p. 37-49

[12] Winterbottom, Alan "Lower Confidence Limits for Series System Reliability from Binomial Subsystem Data" JASA 69 (1974) p. 782-788.